National Cancer Institute

Frederick National Laboratory Advisory Committee

NCI Task Force to Evaluate the NCI/DOE Collaboration

Evaluation Report

October 14, 2020

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Since the beginning of the NCI/DOE Collaboration in 2016, NCI has made a considerable investment the Joint Design of Advanced Computing Solutions for Cancer (JDACS4C) which was established to explore the use of high-performance computing (HPC), machine learning and artificial intelligence (ML/AI), and advanced analytical algorithms uniquely available in the DOE to address critical problems in cancer research. JDASC4C consists of three pilot projects, Uncertainty Quantification (UQ), CANcer Distributed Learning Environment (CANDLE) and ATOM programs. Due to the scope of the evaluation and time constraints this NCI Task Force to Evaluate the NCI/DOE Collaboration (TF) only reviewed the progress by each pilot, as well as the overall NCI management of the collaboration. The TF did not have the opportunity to evaluate the other JDACS4C components in detail. The first four years of the collaboration have identified both strengths and weaknesses for each of the pilot projects specifically, and the NCI management overall. The TF recommends continuing the collaboration, with specific recommendations for each pilot, outreach and engagement, as well as recommendations for the NCI overall management of the collaboration. Briefly:

- Reassess the current pilot projects to determine whether they are appropriately suited to the DOE capabilities, and ensure better alignment and integration with the NCI research community
- Conclude the activities of **pilot 1** and make the aggregated data set broadly available
- Focus **Pilot 2** on refining the coarse-grain models based on data from atomic-level detailed simulations, and experimental validation of model predictions
- Focus the immediate activities of **pilot 3** on decreasing the reporting lag of the SEER registry data by implementing the developed APIs into the SEER registry workflow
- Institute a rigorous approach to the planning of future projects, including establishing an external advisory group for the oversight of each project to ensure greater accountability for scope, productivity and engagement with the NCI research community
- Substantially increase and improve the engagement with the NCI extramural community
- Implement stronger NCI scientific management and oversight, including embedding scientific management into the NCI divisions.

INTRODUCTION

The NCI-DOE partnership is designed to expertise and computational resources uniquely available in the DOE laboratories to push the frontiers of high-performance computing (HPC) and to develop applications that address NCI's pressing need to improve the understanding of cancer biology and develop more effective cancer treatment strategies. The overarching goal of the collaboration is the development of a shared, high performance computing (HPC) ecosystem and advanced analytical algorithms to bring a new class of computing capabilities to bear on cancer research. Specifically, to address the polygenic and architectural complexity of cancer, as well as the extreme heterogeneity therein. This partnership is based on the hypothesis that improving cancer outcomes will benefit from HPC capabilities and algorithms, computational expertise and large-scale data management capabilities that are uniquely available in DOE laboratories. This includes methodology for integrative big-data analysis, machine learning, pattern recognition modeling of complex systems, and predictive simulations at scale.

In June of 2016 the NCI and DOE established a collaboration through a 5-year MOU, as part of the implementation of the recommendations from the Cancer Moonshot Blue Ribbon Panel and the Precision Medicine Initiative. This collaboration, designated the Joint Design of Advanced Computing Solutions for Cancer (JDACS4C), was designed as bi-directional and iterative, ultimately aimed at developing advanced computational solutions for specific areas in cancer research, accelerating hypothesis generation, developing new biological insight and informing the design of next generation high performance computers in an active learning process. The JDACS4C was designed to integrate both data and science in new ways and to create synergy with cancer research and computational science.

The three JDASC4C pilot projects in scope for this evaluation are:

1. A *cellular* level pilot to develop promising new treatment options through predictive computational models of preclinical therapeutic responses
2. A *molecular* level pilot to deepen the understanding of cancer biology using molecular, functional, and structural data to model and characterize RAS membrane biology
3. A *population* level pilot for integrating, analyzing and modeling for precision cancer surveillance

The NCI Task Force to Evaluate the NCI/DOE Collaboration, an independent group of experts, was charged to evaluate whether NCI efforts met the program goals and the investment was commensurate with the outcomes. The task force (TF) was chaired by Dr. Joe Gray from the Oregon Health & Science University, and the full roster can be found in Appendix A. The main objective of this task force was to provide an in-depth technical review and assess the merits of the individual projects, the NCI management of the overall program, and to develop a set of recommendations on whether or how this collaboration should evolve beyond the fifth year. This report is intended to assist NCI leadership in making a final determination about the path forward for the NCI/DOE Collaboration.

The TF evaluated the collaboration by addressing five questions for the overall collaboration, as well as for the individual Pilot projects.

1. What impact has the collaboration overall, and the pilots specifically, had on the cancer research community?
2. How have the unique DOE HPC capabilities and expertise contributed to cancer research?
3. Has the effort effectively engaged the greater cancer research community, and have they benefitted?
4. Are there additional research opportunities for collaboration with DOE and HPC in cancer?
5. Has the NCI oversight been adequate, and should NCI continue to support this collaboration?

The TF developed its recommendations during the course of five virtual meetings between July and October 2020 and through review of presentations, progress reports, and associated materials provided by each of the pilot projects. Dr. Emily Greenspan, the NCI programmatic lead, provided information to the TF spanning diverse measures, including previous presentations, progress reports, FNLAC NCI/DOE Collaboration Working Group materials, financial information, and publications, including bibliometric measures. The NCI pilot leads Drs. James Doroshow and Yvonne Evrard (Pilot 1), Dwight Nissley (Pilot 2), Lynne Penberthy (Pilot 3), and Eric Stahlberg (outreach) provided more detailed information about each of the pilots and highlighted scientific accomplishments and impacts on the broader cancer research community. TF members used the provided information, including their conversations with the pilot leads for the evaluation presented in this report. In addition to providing an overall evaluation of the NCI/DOE collaboration, the TF also evaluated each pilot individually.

Summary of Conclusions

- *Continuing the collaboration*: The TF agrees that the NCI/DOE collaboration is uniquely suited to address critical questions in cancer research that would otherwise be difficult to explore and should continue. There are many lessons learned from the first four years of the collaboration, and the TF recommends that these should be appropriately addressed before moving forward.
- *Reassessment of current pilot projects and adjustment of the level of funding:* TF members were in agreement that the current pilots should no longer be referred to as pilots, rather they are large full-scale projects with budgets exceeding those of other NCI-funded efforts and should thus be evaluated and reassessed as such. As part of the reassessment, the level of funding for each current should be revisited and evaluated.
- *Pilot specific assessments:*
  - **Pilot 1**: The TF recommends concluding this pilot, which failed to meet the foundational premises of this collaboration and lacked integration with similar efforts within the NCI community. This project was initiated prematurely, and there was a lack of appropriate data and insufficient integration with the substantial NCI extramural predictive modeling community.
  - **Pilot 2:** The TF recommends continuing this pilot with future efforts focus on refining the coarse-grain models based on data from atomic-level detailed simulations and on the experimental validation of predictions arising from such models. This was deemed the only project that potentially requires DOE HPC capabilities.
  - **Pilot 3:** The TF recommends continuing this pilot with focus placed on the implementation and multi-institutional deployment of the developed APIs, and engaging

stakeholders beyond the SEER cancer registry community. This should begin with the development of a clear vision. Scientific opportunities include real-time integration of molecular and non-molecular data for diagnostic or therapeutic intervention, and therapeutic decision making and targeting. The close collaboration between the NCI and DOE scientists was critical to the success of this project to date. This pilot leveraged the DOE computational science expertise, rather than the unique HPC capabilities, which were not were not required for these efforts.

- *Selection of future projects:* The TF recommends a more structured and rigorous approach to the development, design, and review of future collaborative projects, including engaging the broader computational and cancer research communities. This would allow for increased awareness and collaboration with the extramural community and leveraging other current NCI investments in systems and computational biology. Most critically, these projects should be perceived and evaluated on the same basis as other extramural NCI-funded research efforts.
- *Increase engagement with the NCI extramural community:* Each pilot project has been largely driven by DOE, and the research communications have been relatively insular to each team. The TF observed that there was a broader reach into the computational science community compared to the cancer research community. They strongly recommend increased outreach and engagement, as well as transparency to the NCI extramural community about the research and progress of the NCI/DOE collaborative projects.
- *Project specific advisory groups*: There should be increased scientific oversight and engagement by the NCI for each project. Much of the drive for the current projects has been from the DOE collaborators with insufficient engagement of relevant extramural cancer research communities. Indeed, the vast majority of publications arising from this collaboration have been led by DOE, with only a minor role played by the NCI. The TF recommends increasing scientific oversight and awareness to the cancer research community by establishing scientific advisory groups comprised of top experts in cancer research for each of the projects funded by this collaboration.
- *NCI Management and Oversight:* To date, the pilots have been insular in their management and self-assessment. The TF recommends that the programmatic and administrative management be handled together and that the scientific management should be moved to the appropriate NCI program staff. For example, the TF thought that the integration of Pilot 3 within the larger SEER program provided strong engagement from NCI, though there was limited engagement of the extramural cancer research community. Pilot 2 benefitted being part of the larger NCI RAS Program, however there was limited ongoing scientific engagement from NCI or the broader RAS biology community. Pilot 1 lacked appropriate oversight, engagement, and involvement from NCI. NCI has been as successful in fostering relationships between diverse scientific communities (for example through the SPORE, PSOC SSBC programs), and for the future success of this collaboration, it is essential to effectively link DOE computational scientists with cancer biologists. Stronger integration of cancer researchers in the ongoing conduct of these pilots could have resulted in findings more relevant to the cancer biology community.

Panel Response to Charge

Pilot 1: Drug Response and Prediction for Preclinical Screening

Question 1: What impact has the collaboration had on the Cancer research Community?

The TF agrees that the question posed for this Pilot is an important one, to make validated efficacy predictions for most drug classes in vivo using cell-based assays or animal models. The experimental goals were (1) develop preclinical drug screening models, (2) define what data scale and properties are needed to optimize predictive oncology experimental design, (3) generate hypotheses regarding method of action and drug or tumor characteristics, and (4) develop effective models for drug discovery. TF members were concerned that there had been little transparency and engagement of the NCI extramural community around this pilot, in particular since this is an active area of research. The primary limitation is the availability of appropriate data, as was the case with this pilot.

It appears as if the research team approached this project from purely a deep learning perspective, without the direct use of underlying biological constraints and knowledge using data that were not expressly generated to support predictive model development and without biological collaborations needed to independently validate the predictive models. While the integration of deep-learning with biological, model-driven methodologies would be extremely valuable, it was felt that the current team lacked appropriate critical biologic expertise to understand the data to provide meaningful insights beyond those that have already been generated based on the available data. The TF highlighted critical shortcomings in the apparent disregard of existing methodologies and approaches to these complex research questions without requiring DOE-level HPC, likely due to the lack of appropriate ongoing engagement from the NCI and the extramural expertise in this area. These findings are reflected by a minimal publication record.

While the predictions from the models showed some level of statistical significance, the TF questioned whether the metrics used were actually biologically relevant and whether he results were meaningful to be able to guide/prioritize drug selection in clinical studies and trials. This disconnect with biologic relevance was even more obvious in the attempts to predict synergistic drug combinations. The impact of these findings on the cancer research community was low and not commensurate with the level of investment. It is the opinion of the TF that greater emphasis on data generation, curation, validation, and distribution to the intra-, and extramural communities would have been more effective in achieving the stated goals and would have provided greater benefit to the cancer research community.

Question 2: How have the unique DOE HPC capabilities and expertise contributed to cancer research?

The team applied deep learning and large-scale computing to the available data. However, the TF agreed that the scale of the available data did not remotely meet the threshold to leverage the unique DOE HPC resources. The approaches used were disconnected from other data modalities that have proven essential in models to predict of drug sensitivity. As a result, this effort did not produce meaningful results, as confirmed by the publication record. It appeared that when this Pilot was developed, there was an incorrect assumption about the availability, quality, and quantity of NCI data for deep learning, as well as perhaps an overstatement of what deep-learning can accomplish, in isolation, in a multi-variate context that cannot be effectively of easily decomposed into smaller and more tractable problems. Decomposition of complex datasets into relatively independent subsets is critical for the success of pure deep learning methodologies. This was unfortunately not the case with this pilot, where decomposition should have relied on additional data sources such as protein-protein/transcriptional interactions, as well as perturbational assays that monitor the molecular response of a biological cell to a variety of exogenous perturbations.

From the information provided to the TF it also appeared that the NCI data were not immediately available upon the start of the collaboration, nor as extensive as anticipated by the modelers. NCI biologists may not have understood the data requirement for HPC well enough to comprehend the scale of data needed for the unique DOE HPC capabilities. This pilot could not succeed without appropriate amounts of data for model training and without the appropriate integration of data modalities, machine-learning, and model-based methodologies. The DOE collaborators moved forward appropriately given the data availability and their methods. To supplement the NCI data, publicly available data sets (as well as data available through other sources) were used for the model simulations. The TF members acknowledged the large effort that went into aggregating data from disparate sources and the model development. Deep learning approaches were applied but given the modest complexity and size of the data, but these did not require the use of HPC resources unique to DOE. That said, full engagement of the NCI predictive modeling community might have yielded a volume of data could justify use of these unique DOE HPC capabilities.

This pilot relied heavily on the DOE collaborators and their computational science expertise and the TF acknowledges that it takes time to develop mutual understanding among people with different areas of expertise, such as biologists and computational scientists. This evaluation should not detract from the deep expertise in machine learning of the DOE collaborators, whose knowledge and scientific leadership is not under discussion. The TF expressed that the DOE collaborators did the best they could, given the circumstances. Rather, it suggests that they were put in an untenable situation where the goals of the project could not be effectively achieved given the lack of appropriate ongoing engagement from NCI to provide biologic expertise, and the amount and type of available data compounded by the heterogeneity of the data. TF members commented that the only successful computational modeling approaches they were aware of in this space, relied on data generation efforts specifically to support the development of the model. This pilot relied on using data already generated by large-scale studies at the NCI, and no additional experiments were designed to generate the necessary data.

Question 3: Has the effort Effectively Engaged the Greater Cancer Research Community, and have they Benefitted?

Previously the FNLAC NCI/DOE Collaborations Working Group had encouraged collaboration among researchers from different disciplines through the NCI Integrated Cancer Biology Program (ICBP) and the Cancer Systems Biology Consortium (CSBC). The JDACS4C Pilot program provided an opportunity to create a hybrid discipline between biology and computer science, but pilot 1 failed to engage with these existing communities and integrate them appropriately into their activities. In fact, the lack of engagement with other NCI-funded investigators involved in predicting drug sensitivity for inclusion of their methodologies and expertise is notable.

The TF commented that it is possible that the data used in pilot 1 had been analyzed by so many groups that not much more could be learned from them. Indeed, a critical concern is that the kind of data that was provided to Project 1 scientists is simply not amenable to effectively predict response to treatment, especially in the context of combination therapy. Comparing the NCI/DOE developed models with other published models and publishing those performance characteristics was a missed opportunity. The cancer research community may have engaged more with the work of this pilot if those types of studies were available. It is unclear how widely, if at all, the developed algorithms have penetrated in the cancer research community.

Question 4: Are there Additional Research Opportunities for Collaboration with the DOE and for HPC in Cancer?

The TF commented that any future efforts for predictive models for preclinical screening should consider the biological models that have been shown to significantly improve the accuracy and sensitivity of predictive algorithms, such as regulatory and signaling network models. The inclusion of biological prior knowledge represents a critical missing component and opportunity for engagement with the extramural NCI community of pilot 1. Examples of biologic variables include the molecular structure of the drugs and targets, the molecular-level response of cells to exogenous perturbations—both genetic (e.g., CRISPR/dCas9-mediated) and drug-related, including pre- and post-treatment data from in vivo models—intratumoral and intrastromal heterogeneity, stromal responses, single cell data, omics data, etc. Indeed, one of the more critical aspects underlying the diversity drug responses in cancer patients is related to the intrinsic molecular and cellular heterogeneity of the disease.

Expertise within the extramural community could be leveraged to develop modular models that could be combined into larger models that would begin to approximate the functions of a living cell. The TF suggested that the combination of multiple models—at different granularity levels, ranging from kinetic to statistical models—into a large simulation framework would be a far more appropriate and exciting effort to take advantage of the unique DOE HPC capabilities.

During the course of this pilot a large repository of aggregated available data was created. Such a repository would be a valuable resource for the cancer community, and the TF recommends those data be made public.

Question 5: Has the NCI Oversight been Adequate and should NCI Continue to Support this Pilot?

The experience with pilot 1 highlighted the need for careful planning and ongoing oversight of these complex computational research projects. The TF recommends that any future projects include near-term milestones for success that can indicate whether a project should continue or be terminated. It is unclear what the role of the NCI has been with this Pilot, aside from providing data. The TF contends that the lack of scientific oversight and engagement throughout the lifecycle of this pilot was a failure of NCI management. Stronger involvement from NCI, particularly from the extramural divisions, may have enabled more engagement from the cancer research community, greater biologic relevance, and increased the reach of the publications. Additionally, the lack of an external scientific advisory board was noted as a critical issue. The TF appreciates that this is a complex problem, but that the HPC approaches using the currently available data were premature. Rather, it is the sentiment of the TF was that this effort would have benefited from attention to data generation, curation, and scientific understanding of the problem before launching the computational effort. The TF recommends concluding the current activities of this pilot.

Pilot 1 Summary:

- This pilot should be concluded
- Increased efforts should have been placed on experimental data curation and validation
- This pilot had a laudable vision but was initiated prematurely, it lacked the appropriate data and mechanistic understanding to justify the level of HPC computing
- Pilot lacked appropriate and ongoing NCI management, oversight, and engagement
- Data aggregation was a large effort on the part of DOE and the TF recommends those data be made available to the cancer research community

Pilot 2: Improving Outcomes for RAS-Related Cancers

Question 1: What impact has the collaboration had on the cancer research community?

This pilot benefitted from being integrated into the activities of the existing RAS Project at the FNLCR. The RAS Project has had considerable interest and oversight from NCI leadership and the inclusion of the DOE collaboration was a natural fit. It is due to this integration that pilot 2 into has had some (although insufficient) visibility in the broader RAS biology community. Members of the TF agreed that among the NCI/DOE collaborations, Pilot 2 has had been the most successful in leveraging the DOE HPC capabilities. The computational modeling simulations from this pilot are hypothesis-generating, exploring possible mechanisms for where and how RAS/RAF can engage, and the results from these models have the potential to impact the biophysics and RAS biology community broadly. This collaboration has brought new insights into the ability to drugging RAS, for example the identification of an additional potential kinetic state of RAS. However, the TF noted that to date the impact of this work has been mainly focused on the computational and biophysics communities, and penetration into the general RAS/ERK signaling community, as well as experimental validation of the computational models, have been limited.

Based on the citation information from the publications, it appears that the work from this pilot has had limited impact. As with the other pilots, the TF noted that the majority of the publications were focused on computational advances, rather than the biological insights provided by the models. A minority of the publications show co-authorship between the main RAS biology group at the FNLCR and the DOE collaborators, therefore it is difficult to discern the level of interaction between the collaborators. A number of the publications that were provided as outputs of this pilot are on topics are somewhat removed from RAS/membrane studies (or in general from RAS/RAF signaling), the primary goal of the project. The finding of an additional potential kinetic state of RAS bound to the membrane was significant. Other more recent contributions (Travers et al., Biophys. J 2020[1], Neal and Garcia, Biophys J, 2020[2]) suggest that the RAS-binding domain (RBD) and cysteine-rich domain (CRD) in RAF act in a concerted fashion to drive RAF activation, and that the PM has dual action as an allosteric activator and inhibitor of RAS. However, it is notable that only one of these publications includes NCI authors. TF members were surprised that some of the RAS Program staff at FNLCR, world leaders in RAS biology, do not appear to be intimately involved with this collaborative project based on publication authorship.

Overall, the TF felt that the productivity of this pilot was good, and significant advances have been made in terms of algorithm development. However, the simulations themselves have yet to yield significant new insights that could not be obtained (and in some cases *were* obtained) using more conventional resources and approaches. Additional emphasis should be placed on further integrating the broader RAS biology community (both at the FNLCR and the extramural program) into the activities of this pilot, as well as experimentally validating the computational model findings.

Question 2: How have the unique DOE HPC capabilities and expertise contributed to cancer research?

The TF agreed that the collaboration with DOE is essential for the premise of pilot 2, and an excellent application of the unique HPC capabilities and expertise from DOE scientists. Two DOE National Laboratories contributed their computational capabilities for this collaboration. While the goals of the pilot justified the use of HPC computing, and the research has the potential to provide unique information to the RAS biophysics community, the simulations themselves have yet to yield significant new insights. Specialized hardware architecture, such as FPGA chips, may enable solving the same problems with less computing power – concepts that may influence the design of future DOE HPCs. TF members recommend developing techniques toward refining their coarse-grained models based on data from atomically detailed simulations to further explore protein/protein interactions and interfaces.

[1] Travers, T., López, C.A., Agamasu, C., Hettige, J.J., Messing, S., García, A.E., Stephen, A.G., Gnanakaran, S., 2020. Anionic Lipids Impact RAS-Binding Site Accessibility and Membrane Binding Affinity of CRAF RBD-CRD. Biophysical Journal. doi:10.1016/j.bpj.2020.06.021

[2] Neale, C., García, A.E., 2020. The Plasma Membrane as a Competitive Inhibitor and Positive Allosteric Modulator of KRas4B Signaling. Biophysical Journal. doi:10.1016/j.bpj.2019.12.039

The TF also strongly recommends that more, and interactive, biological/biochemical experimental validation be performed to test the predictions of the DOE models and to create additional hypotheses.

Question 3: Has the effort effectively engaged the greater cancer research community, and have they benefitted?

Due to the integration with the existing RAS Initiative, pilot 2 was able to leverage those distribution channels to more broadly reach the RAS research community. Communications at large national cancer meetings and during RAS Initiative meetings have, to some extent, helped to increase awareness of the NCI/DOE RAS modeling efforts. Yet despite general awareness of the pilot by the RAS cancer research community, it is unclear that the cancer community is productively engaged with the work of the pilot. For example, while the Di Natale[3] publication won best paper at SC19 (The International Conference for High Performance Computing, Networking, Storage, and Analysis), there does not seem to be broad awareness of the work in the general RAS biology community. TF members and their colleagues noted that while they heard of the publication at an FNLCR RAS meeting, they did not have any real awareness of the research beyond that. Another member went on to say that any knowledge of the research prior to publication was through personal interactions, rather than community engagement from the NCI/DOE Collaborations or RAS Initiative teams. Comments such as these are likely indicative that the access to the information generated by this collaboration by the RAS biological community needs to be improved. The majority the macroscale like imaging research is highly specific and technical and not very accessible to the larger biological community that is working on RAS. TF members recommended that the modelers should strive to make the biological and conceptual implications of their work more generally accessible to the RAS community to provoke new experiments. To more broadly engage the general RAS community, the TF recommends that the NCI and DOE teams prepare a publication specifically for the cancer research community in a well-known journal on the implications of computer-based simulations for RAS/RAF pathway signaling written in language accessible to the cancer research and signal transduction community.

It would also be of value to compare the performance of the NCI/DOE model simulations to those that have been developed on other platforms. Emphasizing the utility of the DOE developed algorithms in this space, as well as the use of the UQ. This might lead to broader dissemination and use of the developed tools. TF members also recommend that members of the RAS community who perform modeling and computational experiments be better engaged, possibly by enabling access to the unique resources provided by the DOE collaboration (e.g., via a model/construct more like other big NCI projects such as TCGA or ENCODE).

Question 4: Are there additional research opportunities for collaboration with DOE and HPC in cancer?

---

[3] Di Natale, F., Bhatia, H., Carpenter, T.S., Neale, C., Kokkila-Schumacher, S., Oppelstrup, T., Stanton, L., Zhang, X., Sundram, S., Scogland, T.R.W., Dharuman, G., Surh, M.P., Yang, Y., Misale, C., Schneidenbach, L., Costa, C., Kim, C., D'Amora, B., Gnanakaran, S., Nissley, D.V., Streitz, F., Lightstone, F.C., Bremer, P.-T., Glosli, J.N., Ingólfsson, H.I., 2019. A massively parallel infrastructure for adaptive multiscale simulations, in: doi:10.1145/3295500.3356197

There are many opportunities to expand the current work with RAS, both biologically and computationally. Experimentally validating computational models will further strengthen the existing simulations. There is a lot of interest in studying protein-lipid interactions in atomic detail, and the collaboration could move in that direction, thereby making more effective use of the DOE HPC power. The molecular modeling work could also be expanded to other proteins, for example HRAS or other molecules that are commonly mutated in cancers. Additionally, the TF recommended further exploring the use of the UQ project in quantifying biases in the model simulations when bridging micro-to-macro scales.

Question 5: Has the NCI oversight been adequate and should NCI continue to support this Pilot?

The work of this pilot was integrated into an existing NCI program with direct oversight from NCI FNLCR leadership. The existing infrastructure enabled broader outreach to the RAS community, including presentations are large cancer conferences and RAS-specific conferences. Despite this integration and oversight, it is unclear how well the RAS community, beyond the biophysics community, was engaged. In addition to the engagement from the FNLCR, there should have been broader programmatic engagement with the NCI extramural divisions. Through this partnership with DOE, NCI researchers could have access to level a of computation is truly unique and not accessible in other ways. To further increase the impact and value of these simulations to provide new structural insights to the RAS community, the TF strongly recommends closer integration with the biological RAS expertise with the NCI divisions and the extramural community. Such integration could occur via collaborative projects with the FNLCR, or by providing extramural funding to grantees in support of an NCI/DOE project. The TF also recommends using the expertise of the FNLAC RAS Working Group as the scientific advisory board for this pilot to provide oversight and guidance to the pilot scientists, with the goal of enabling scientific breakthroughs rather than incremental progress.

Pilot 2 Summary:

- The TF recommends continued efforts in support of protein model simulations as emphasized by pilot 2
- This project is uniquely suited to leverage the DOE HPC capabilities, productivity has been acceptable, and the TF recommends continued support for this effort
- Experimental validation computational model findings should be a cornerstone of the collaboration
- Substantially improve outreach and engagement to the broader RAS biology and molecular simulation communities
    - Engagement with the biophysics community is good but can be improved
- Develop a strategy to engage NCI extramural researchers into direct collaborations with the NCI and DOE, for example by providing funding opportunities to collaborate with the NCI and DOE investigators
- Leverage the FNLAC RAS Working Group for scientific leadership and oversight of this project

PILOT 3: POPULATION INFORMATION INTEGRATION, ANALYSIS, AND MODELING FOR PRECISION SURVEILLANCE

QUESTION 1: WHAT IMPACT HAS THE COLLABORATION HAD ON THE CANCER RESEARCH COMMUNITY?

The focus of this pilot has been on working with the NCI Surveillance, Epidemiology, and End Results (SEER) program that supports research on the diagnosis, treatment, and outcomes of cancer, including reporting on U.S. cancer incidence and mortality. Currently, manual data curation results in a 2-year delay in reporting these data to the public, and through this effort SEER is working to enhance the reporting to near real-time. While the TF determined that this pilot did not depend on DOE HPC capabilities, the combination of DOE computation and registry science will indeed be very impactful if it improves the availability of SEER data in near real-time. TF members were unanimous in their agreement that the successful implementation of ML/AI into the cancer registries would have a big impact on both the cancer research and cancer clinical communities.

The strengths of this project are based on the large volumes of data that the SEER program was able to share via this pilot, as well as the close collaboration between the NCI and DOE pilot teams. The TF acknowledges the significant achievement in working with the individual registries to obtain legal permissions to share their data. The focus this project has been on using ML/AI to automate data extraction of five key elements from pathology reports for SEER reporting. The algorithm has achieved the goal of 97% accuracy, roughly the accuracy of a human annotator.

The TF noted that the ML algorithm for extracting the five key elements is relatively simple and aimed mostly to the central cancer registry and SEER registry community but may not be as useful for hospital and community-based cancer registries that serve NCI designated cancer centers. While the current five elements that are automatically extract are useful, their utility to the broader community is limited because the lack of information on the stage of disease, and other critical outcomes (for example treatment, co-morbidities, and biomarkers), that are critical for case ascertainment for therapeutic targeting. TF members acknowledged that stage of disease is difficult extract because the data is in multiple places/records. They also noted that the privacy preserving method was relatively simple (removal of terms that obviously contain personal information from publicly available sources and the algorithm was subsequently retrained on the cleaned data), but that no demonstrable proof of privacy was offered. They also raised the concern that there does not appear to be a clear path for wider implementation. Specifically, to facilitate the integration and implementation of the application programming interfaces (APIs) into the clinical environment of the three critical target users: 1) SEER registries, 2) Central Cancer Registries and 3) hospital and community based cancer registries. Moreover, it was unclear to the TF members that the current team has the appropriate expertise to facilitate the integration and clinical implementation. Future efforts should include clear implementation plans for integrating these algorithms into the registry workflow.

The lack of standard data definitions for common elements is a major barrier in the field of machine learning and it limits the ability to analyze and automate data extraction from clinical documents. Currently, there are too many diagnostic codes for stage of disease, etc. The TF recommends that the cancer community should come together to harmonize definitions to further advance ML/AI capabilities and application in cancer research and care. More importantly, a clear focus needs to be placed on developing a more comprehensive data model for computationally enabling the SEER registries, and well as the broader cancer registry (Central Cancer Registries and Hospital/Community base cancer registries) community.

Question 2: How have the unique DOE HPC capabilities and expertise contributed to cancer research?

For pilot 3 the DOE computational expertise has been critical, as emphasized by the pilot's publications. Though HPC was used throughout this pilot to explore different types of models and for algorithm development, HPC at a scale unique to DOE was unnecessary. The TF thought that this work illustrated a machine learning problem, and the DOE computational infrastructure was leveraged to enable rapid development, testing, scaling, and deployment of models and APIs. The DOE collaboration enabled these APIs to be hardened and dockerized for broad use by the SEER cancer registry community. However, the TF thought that not much consideration was given to the feasibility of implementing these APIs in Central Cancer Registries or hospital/community-based registries.

A major strength of the collaboration was that the SEER program was able to provide large volumes of structured data, critical for training for the computational models. Currently, those data are not available to the outside research community due to privacy concerns from personally identifiable information (PII). Every effort should be made to address the privacy concerns and broadly share these data with the extramural research community, in particular since this effort was funded by the Cancer Moonshot Initiative and data sharing is a key objective. If these same data were available to the extramural community much of the research might have been done in academic or commercial settings.

Question 3: Has the effort effectively engaged the greater cancer research community, and have they benefitted?

The APIs that were developed through this collaboration are broadly available to the community, but their utility is mainly focused in the SEER cancer registry community, including the API for real-time automated extraction of five key data elements from pathology reports, privacy preserving API, disease recurrence API, and the reportability API. Overall, the TF thought that the data models are focused solely on the SEER and registry community, rather than the Central Cancer Registries, hospital/community-based registries, or broader cancer research community.

The TF noted that availability of appropriate, PII-free data has been the limiting factor in the community for the development of ML/AI approaches. By making the data used to train these algorithms available to the broader modeling community the impact of the pilot on the greater cancer research community could be dramatically improved. The TF recommends developing a mechanism that would allow the research community access to analyze these data, including PII. For example, by shipping a dockerized container with an algorithm to the SEER data. The TF recommends that these capabilities be further extended to the larger cancer registry community and beyond.

The work from this pilot resulted in the largest number of publications compared to the other pilots. To more broadly reach the ML/AI and cancer research communities the TF recommends convening a national meeting to gain input from commercial entities and subject-matter experts on future directions for the pilot.

Question 4: Are there additional research opportunities for collaboration with DOE HPC in cancer?

There are many uses for the automated coding of medical information, including imaging and treatment data. The TF noted that the current algorithms would have greater utility if it would be possible to make an automatic determination on the stage of disease based on the pathology reports. Other communities recommended for closer collaboration in the future include the radiology community for the incorporation of pathology imaging and the radiation oncology community for dose and fractionation data. This pilot uniquely is poised to scale its work to other NCI programs like the Information Technology for Cancer Research (ITCR) program and integrate into the broader cancer research modeling community to increasing the possibility for breakthrough cancer research. Increased emphasis should be placed on coordinating this effort with cancer registries at NCI-Designated Cancer Centers and with the new Childhood Cancer Data Initiative (CCDI).

Question 5: Has the NCI oversight been adequate and should NCI continue to support this Pilot?

The close interaction between the NCI and DOE investigators has been critical to the success of this pilot, in particular because the unprecedented access to the cancer registry community. NCI staff has been very engaged throughout the entire process, developing close and ongoing dialogues with their DOE colleagues. The integration with the SEER program has been beneficial to the outcome of this pilot. However, an additional science officer in a different NCI division could have enabled broader engagement from the NCI computational science community.

From the information that was provided to the TF, members noted that the goals for this pilot have shifted over time. From automating registry data extraction and decreasing the time to reporting, to supporting clinical researchers in identifying patients for trials, to linking registries to other databases to provide a comprehensive view of the cancer patient and their treatments. In the near term the TF recommends focusing efforts on the implementation of the developed APIs into the cancer registry workflow and reducing the SEER reporting time for all registries. For future directions related to population information integration, analysis, and modeling for precision surveillance the TF recommends convening an external advisory group to provide clear and focused directions for moving forward.

Pilot 3 Summary:

- This pilot project should continue with a focus on implementation and broader applicability to the cancer registry community beyond SEER
- The close and ongoing collaboration between the NCI and DOE teams contributed to the success of this pilot

- DOE computational science expertise was critical to this project, rather than HPC
- Future directions should be specified with the help of an external advisory group. This group should include expertise on the implementation of open source software into commercially supported tools, cancer diagnosis, disease stratification, therapeutic targeting, as well as pathologists, radiologists, and medical oncologists or other clinical trialists
- Increased engagement with the NCI extramural community, for example by convening workshops.

OVERALL RECOMMENDATIONS FORM THE TASK FORCE

SHOULD THE NCI CONTINUE TO SUPPORT THE COLLABORATION?

Since the start of the collaboration in 2016, the NCI and DOE have successfully initiated an effort to explore the extent to which the HPC capabilities and computing expertise now resident in DOE National Laboratories, can be brought to bear effectively on problems of importance to the cancer research community. While there were some challenges, the TF thought the overall effort has been sufficiently productive and recommends that the NCI/DOE collaboration should continue with the level of funding adjusted to be appropriate to achieve clearly defined project goals. However, the three initial pilot projects that were part this collaboration have illuminated both strengths and weaknesses related to the selection, outreach, experimental validation, and the ongoing NCI management.

The TF recommends that these weaknesses should be addressed before proceeding with further funding or other collaborative projects. Specifically, the TF determined that pilot 1, drug response prediction for preclinical screening, was launched prematurely without enough forethought given to the scientific question. As a consequence, this pilot was not well integrated with the larger NCI predictive modeling community, and it lacked the well curated data sets needed for ML. The TF recommends concluding this pilot. Pilot 2, improving outcomes for RAS-related cancers, which focusses mainly on modeling RAS dynamics has made use of the unique DOE HPC capabilities and has generated new scientific insights for the well-organized RAS research community. The TF recommends that pilot 2 should continue with an emphasis on more fine-grained modeling, as well as experimental validation of the predictive models. Pilot 3, population information, analysis, and modeling, is strongly focused on the integration of natural language processing (NLP) in support of the NCI SEER program and the cancer registry community. This pilot has accelerated the acquisition of information into the SEER registry but has missed opportunities for broader impact by not engaging with the broader NCI research community. Though the collaboration this pilot was able to leverage the DOE computational expertise. However, Pilot 3 did not appear to require DOE HPC capabilities, one of the initial goals of the collaboration. The TF recommends that this pilot continue after a reassessment of the overall program objective and plans for better integration with the larger NCI clinical cancer research community.

ARE THERE ADDITIONAL RESEARCH OPPORTUNITIES FOR COLLABORATION WITH DOE AND HPC IN CANCER?

The deliberations of the TF were mainly focused on the evaluation of the current pilot projects, but the members did have some specific recommendations for future directions. As illustrated by the success of pilot 2, there are many aspects of molecular dynamic modeling that can benefit from the unique DOE HPC capabilities and modeling expertise. It is likely that the need in this area for HPC will increase as emerging cryoEM tools generate high resolution structures of protein complexes and molecular assemblies thereof. To date this work has focused on the modeling and experimental validation of RAS, and these same methods could be leveraged for other proteins or signaling interactions.

Other opportunities that can leverage this unique partnership include the emergence of well curated, information rich omic and image data sets linked to precisely defined perturbations and biological or clinical responses. These rich data will enable the development of ML/AI strategies to predict biological and clinical outcomes. HPC and DOE expertise in computer science can play a major role in the development of these predictors once the needed data sets are available. However, as illustrated by the findings of Pilot 1, substantial resources will need to be invested in data curation, organization and community engagement in order to appropriately power this area of collaboration.

During one of the program presentations, the development of a cancer patient digital twin for predictive oncology was mentioned as a possible future direction for the collaboration. The development of a cancer digital twin is ambitious, and TF members are not convinced that appropriately curated data sets exist to successfully develop these models. The TF recommends assessing the data availability and feasibility of these types of studies to avoid overfitting the models and having a similar experience as was had with pilot 1. Rather, the TF recommends that initial efforts might focus on developing predictive models of living cells. It would provide a unique opportunity to collaborate with the broader computational cancer research community and leverage the DOE expertise and HPC infrastructure to combine different models from different groups into a more comprehensive model.

Definition of biologically or clinically important features of n-dimensional images is a growing need in cancer research that may benefit from DOE HPC capabilities, data management infrastructure and computational expertise. Hosting and dissemination of these large data sets could be an important part of this collaboration in the future.

IS THE INFORMATION GAINED SUFFICIENTLY VALUABLE OR POTENTIALLY VALUABLE TO JUSTIFY THE EXPENSE?

The success of the pilot projects as measured in terms of productivity, community engagement and high impact findings varies between the projects and the productivity does not justify the expense – so far. The assessment of the TF is that this is due in large measure to the rapid full-scale launch of the pilots without adequate planning or engagement of the NCI extramural community. As a consequence, some important problems worthy of DOE and NCI collaborative attention were likely missed. In the rush to stand up these pilots, important capabilities in some DOE laboratories were not used, data sets that would have been critical to the research were not available, and engagement of the broader NCI community engagement needed for experimental validation of computational findings was missing. It might have been more cost effective if the “pilot” projects were indeed pilots, with smaller budgets requiring more focus and planning, and earlier assessment of their feasibility tied to determinations to continue to develop a larger project or pursue other avenues.

In spite of these shortcomings, the RAS and SEER pilot projects did illustrate the power and promise of this collaboration. This suggests that with proper planning and extramural community engagement the DOE/NCI computational collaboration will enable important studies that would not otherwise be feasible. Execution of such projects could justify the considerable expense.

Are there lessons learned that could improve the overall collaboration?

It is the assessment of the TF that these initial pilot projects were not adequately peer reviewed before they were launched at full scale. It is important to identify projects that are uniquely suited for the collaboration in that they can best be done leveraging capabilities available only at the DOE National Laboratories. For other projects of this scale, there is an extended period of planning and community engagement that did not occur in this case. The JDACS4C Pilots proceeded to the implementation phase without firm goals. Collaboration between the computer science and cancer biology communities has great potential, but these pilots have shown that taking the time for initial concept development, capability matching, and goal definition is critical. HPC technology has progressed rapidly, and many computationally intensive problems can now be addressed successfully using the extensive computational capabilities that are available to the NCI community (e.g. via academic centers, cloud-based infrastructures like AWS, etc.) Alternatively, other specific resources of the DOE, such as non-HPC prowess including computer science expertise, large scale data management and distribution, and multi-institutional project management might also be amenable to collaborative efforts.

The NCI extramural community was not adequately engaged in the selection and ongoing discussions about the conduct and progress of the individual pilot projects. As a consequence, important problems were missed, appropriate data sets were not provided, ongoing NCI supported research programs (e.g. HTAN, CSBC, PSON, CCDI, SPOREs) that might have been mutually beneficial with the DOE/NCI collaboration were not engaged, and computational predictions were not adequately tested experimentally. Well curated data sets needed for computational attention were missing in some cases. It is important to ensure that the high-quality data sets that are needed for these types of projects are available before launching at full scale. NCI resources will need to be allocated in some cases to develop those data sets in preparation for DOE/NCI collaborative projects. Going forward, the problem specifications, data collection and follow-on biological validations should be driven by the NCI cancer research community.

The TF acknowledges that while central programmatic oversight at NCI is appropriate for programmatic and administrative purposes, such as accounting, they feel strongly that the scientific management and accountability is best handled by NCI program staff knowledgeable about the specific area of science and its community. From an NCI scientific management perspective, specific projects would be much better managed in the specific extramural division of the NCI that oversees that area of science. This would allow for easier identification of collaborative or outreach efforts with the appropriate NCI communities. In addition, the TF also recommends establishing an external advisory group for each project to provide ongoing scientific guidance and oversight to ensure greater accountability for scope, productivity and engagement with the NCI research community. In some instances, existing Working Groups, like the FNLAC RAS Working Group should be leveraged for their expertise.

There are a number of projects of high importance to the NCI cancer mission (as described above) that are appropriate for the DOE/NCI collaboration, such as those that leverage DOE expertise in computer science, large scale project management and HPC technology. Development of "worthy" projects takes time and should be done with full engagement of the NCI intramural and extramural communities, in addition to the DOE computational science community. A shortcoming of this collaboration is that the projects have been very insular and lacked engagement from the cancer biology community. Increased communication between cancer biologists and their DOE counterparts could have resulted in greater scientific impact on the cancer research community. There may be opportunities for the development of future projects via planning grant mechanisms that engage both NCI and DOE communities providing a path to develop pilots intended to lead to fully funded DOE/NCI collaboration projects, as well as access to unique DOE resources.

NATIONAL INSTITUTES OF HEALTH
National Cancer Institute
Frederick National Laboratory Advisory Committee
NCI Task Force to Evaluate the NCI/DOE Collaboration

CHAIR

Joe W. Gray, Ph.D.
Professor and Gordon Moore Endowed Chair
Department of Biomedical Engineering
Director, Center for Spatial Systems Biomedicine
Associate Director for Biophysical Oncology
Knight Cancer Institute
Oregon Health and Science University
Portland, Oregon

MEMBERS

Michael John Becich, M.D., Ph.D.
Chairman and Distinguished University Professor, Department of Biomedical Informatics
Professor of Pathology, Information Sciences, Telecommunications and Clinical/Translational Sciences
Director, Center for Commercial Application (CCA) of Healthcare Data
Associate Director for Cancer Institute (UPCI)
Associate Director, Clinical and Translational Science Institute (CTSI)
University of Pittsburgh School of Medicine
Pittsburgh, Pennsylvania

Otis W. Brawley, M.D., M.A.C.P., F.A.S.C.O., F.A.C.E.
Bloomberg Distinguished Professor of Oncology and Epidemiology
The Sidney Kimmel Comprehensive Cancer Center
Johns Hopkins University
Baltimore, Maryland

Andrea Califano, Ph.D.
Clyde and Helen Wu Professor of Chemical and Systems Biology
Director, JP Sulzberger Columbia Genome Center
Associate Director for Bioinformatics
Herbert Irving Comprehensive Cancer Center
Columbia University
New York, New York

Alemayehu (Alex) Gorfe, Ph.D.
Associate Professor
Department of Integrative Biology and Pharmacology
University of Texas McGovern Medical School-Houston
Houston, Texas

Susan Gregurick, Ph.D.
Associate Director for Data Science
Director of NIH Office of Data Science Strategy
National Institutes of Health
Bethesda, Maryland

Robert L. Grossman, Ph.D.
Jim and Karen Frank Director for Data Intensive Science
Frederick H. Rawson Professor of Medicine and Computer Science
Department of Medicine
University of Chicago
Chicago, Illinois

Benjamin G. Neel, M.D., Ph.D.
Professor of Medicine
New York University School of Medicine
Director
Laura and Isaac Perlmutter Cancer Center
New York University Langone Health
New York, New York

Sylvia Katina Plevritis, Ph.D.
Department of Biomedical Data Science
Professor, Departments of Biomedical
Data Science and Radiology
Stanford University School of Medicine
Stanford, California

Eytan Ruppin, M.D., Ph.D.
Chief Cancer Data Science Laboratory
Center for Cancer Research
National Cancer Institute
Bethesda, Maryland

Matthew Trunnell, M.S.
Vice President and Chief Data Officer
McIlwain Family Endowed Chair
Fred Hutchinson Cancer Research Center
Seattle, Washington

Cheryl L. Willman, M.D.
Maurice and Marguerite Liberman
Distinguished Chair in Cancer Research
Director and CEO, University of New Mexico
Comprehensive Cancer Center
University of New Mexico
Albuquerque, New Mexico

Katherine A. Yelick, Ph.D.
Professor
Computer Science Division
Department of Electrical Engineering
and Computer Sciences
University of California, Berkeley
Berkeley, California

Executive Secretary:

Daniel Gallahan, Ph.D.
Director
Division of Cancer Biology
National Cancer Institute
Bethesda, Maryland